Prediction, Belief, and Markets

http://aaaimarketstutorial.pbworks.com

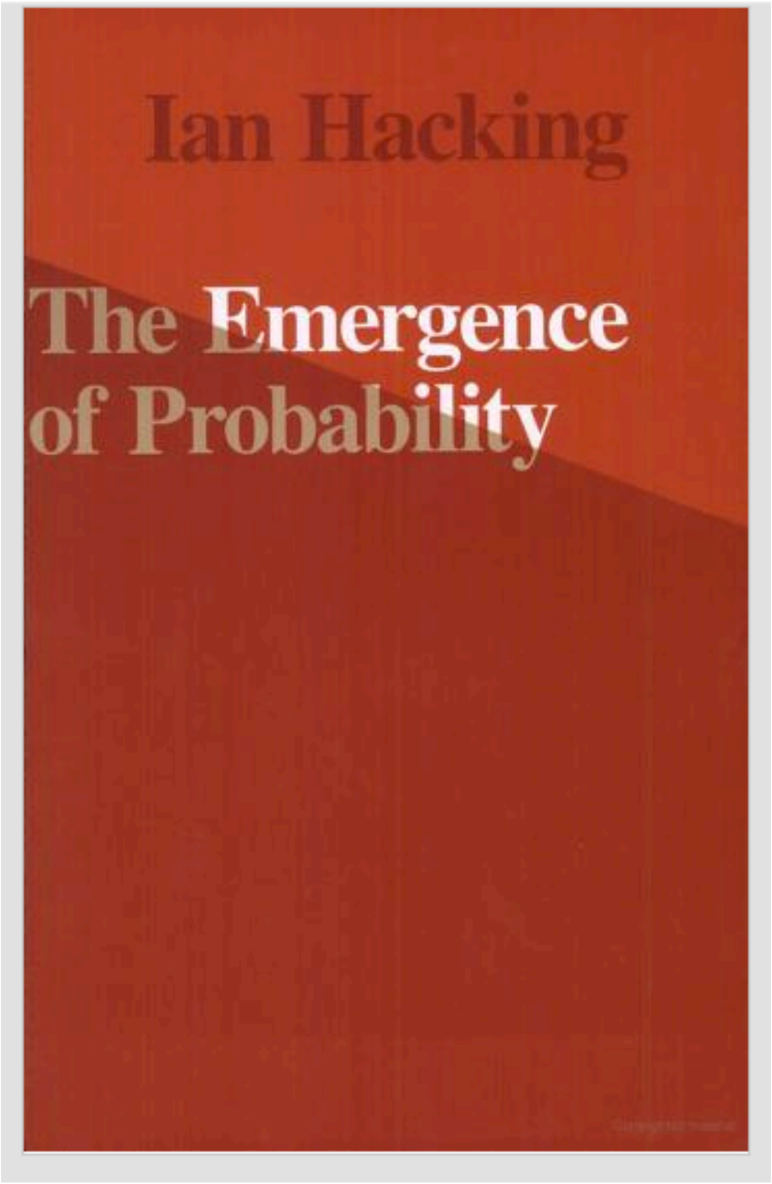
Jake Abernethy, UPenn → UMich
Jenn Wortman Vaughan, MSR NYC

July 14, 2013

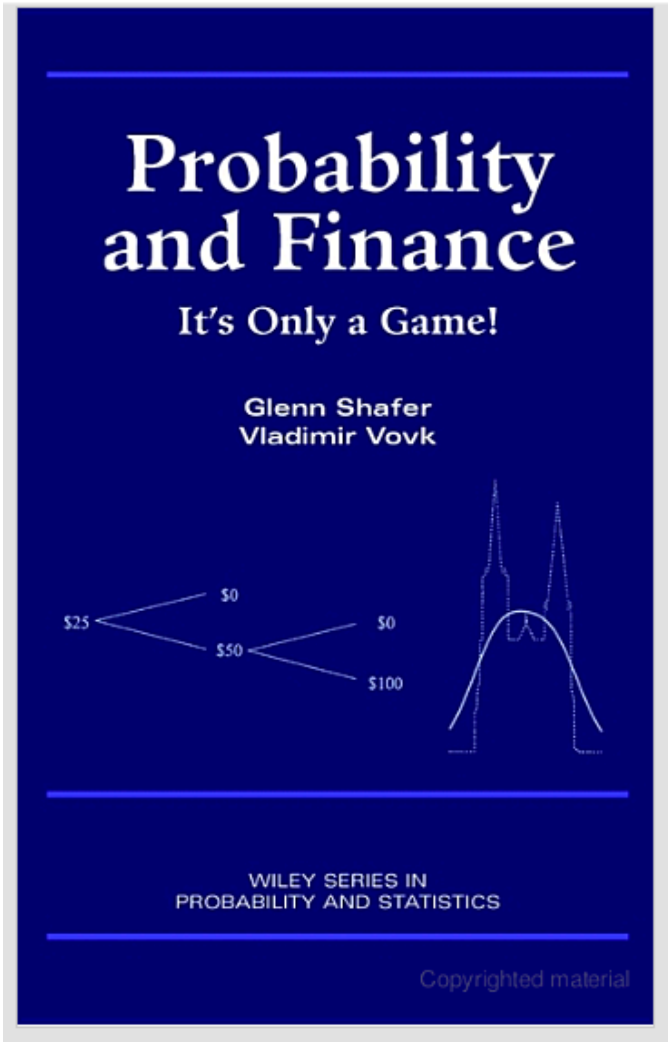
Belief, Prediction, and… Gambling? A Short History Lesson

- The earliest references to probability calculations arose directly through the study of games of chance, like dice.
- We have evidence of human gambling as early as ancient Egypt, more than 3000 years ago.
- We have no evidence of probabilistic analysis until the time Pascal in the late 1600s.

The Emergence of Probability



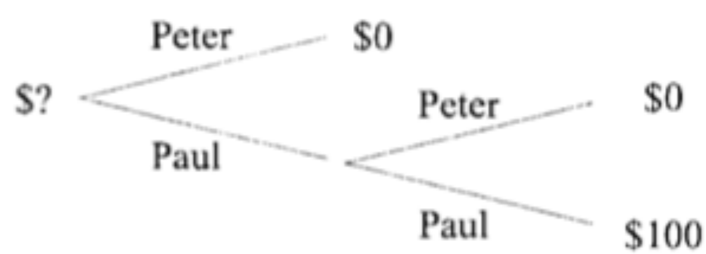
Probability and Finance: It's Only a Game!



Letters b/t Pascal and Fermat, 1654

- The beginning of mathematical probability is often dated to letters between Pierre de Fermat and Blaise Pascal.
- The setting discussed by Pierre and Fermat: two players are playing a game, they need to get *n* points to win, and the winner receives prize.
- How to divide the prize if the game is cut short?

Fig. 2.1 Paul wins $100 if he wins both of the next two rounds.



Bruno de Finetti: "PROBABILITY DOES NOT EXIST"

- De Finetti, well-known in probability theory, had the view that we may only think about probabilities in terms of **rates of betting**.
- That is, the laws of probability can be viewed as resulting from simple "no-arbitrage" conditions on these rates

Betting can be used to elicit beliefs

- Economists, in particular, are very keen on betting with each other. Alex Tabarrok, in marginalrevolution.com: “A Bet is a Tax on Bullshit”
- There’s been some recent debate in the Econ blogosphere about the extent to which bets really do reveal beliefs. Noah Smith, of noahpinion.com: “The mistake is looking at the risk and return of single assets instead of portfolios. Basically, the risk of an asset… is based mainly on how that asset related to *other assets in your portfolio*.”

How do I find someone to bet with? Answer: A **Prediction Market**

- Prediction markets have existed for over 200 years.
- Typically, odds (prices) are set by supply and demand.
- People began to notice: the market prices are generally **very** accurate, and provide better predictors than expert assessments, etc. What's going on?
- Robin Hanson: "Rational expectations theory predicts that, in equilibrium, asset prices will reflect all of the information held by market participants. This theorized *information aggregation property* of prices has lead economists to become increasingly interested in using securities markets to predict future events."

Outline

Before the break:

1. Predictions Markets in Practice
2. Eliciting beliefs with proper scoring rules
3. Bregman divergences + proper scoring rules
4. Hanson' s Market Scoring Rule

After the break:

1. Securities markets
2. Duality & connections to online learning
3. Handling very large outcome spaces
4. Overview of additional topics

The Most (in)Famous Prediction Market

intrade
The World's Leading Prediction Market

- 1999: Intrade founded by John Delaney
- 2003: Acquired by TradeSports in 2003, later splits off after TradeSports closes down in 2008
- 2004: Intrade gains notoriety during Bush/Kerry election for providing continuous forecasts throughout campaign
- May 2011: Founder John Delaney dies at age 42 while climbing Mt. Everest, less than 50 meters from summit
- Nov. 2012: US regulator CFTC files suit against Intrade, leading Intrade to disallow US customers from betting
- Mar. 2013: Due to "financial irregularities", Intrade halts trading, freezes all accounts. Still remains in legal limbo.

Example: Intrade

The Dark Knight Rises to break the all-time opening weekend box-office record

Last prediction was: **$4.08** / share

Today's Change: ▲ **+$0.02** (+0.5%)

Contract Type: **0-100** ?

40.8%
CHANCE

NASA to announce discovery of extraterrestrial life before midnight ET 31 Dec 2012

Last prediction was: **$0.43** / share

Today's Change: -

Contract Type: **0-100** ?

4.3%
CHANCE

Iowa Electronic Markets (IEM): Legal and with Real Money

- Founded in 1988 at the University of Iowa for the purpose of research in market prediction accuracy
- Received a “no action” letter from the CFTC, permitting them to facilitate unregulated betting. (Such letters are apparently “no longer being given out”)
- On the downside, the IEM must obey a certain set of conditions. Most notably, individual traders may deposit no more than $500.

Predictious: A New *Bitcoin*-based Prediction Market

Example: Inkling Markets

Internal prediction markets used within companies

Markets in Practice

Questions:

1. What are different market "mechanisms"?
2. How quickly do markets incorporate information?
3. How accurate are market prices, vis-à-vis prediction?

“Arrow-Debreu” Securities

“Arrow-Debreu Security”: Contract pays \$10 if X happens, \$0 otherwise. If I think that Pr(X) = p then I should:

- Buy this security at any price less than \$10p
- Sell this security at any price greater than \$10p

Current price measures the population’s collective beliefs

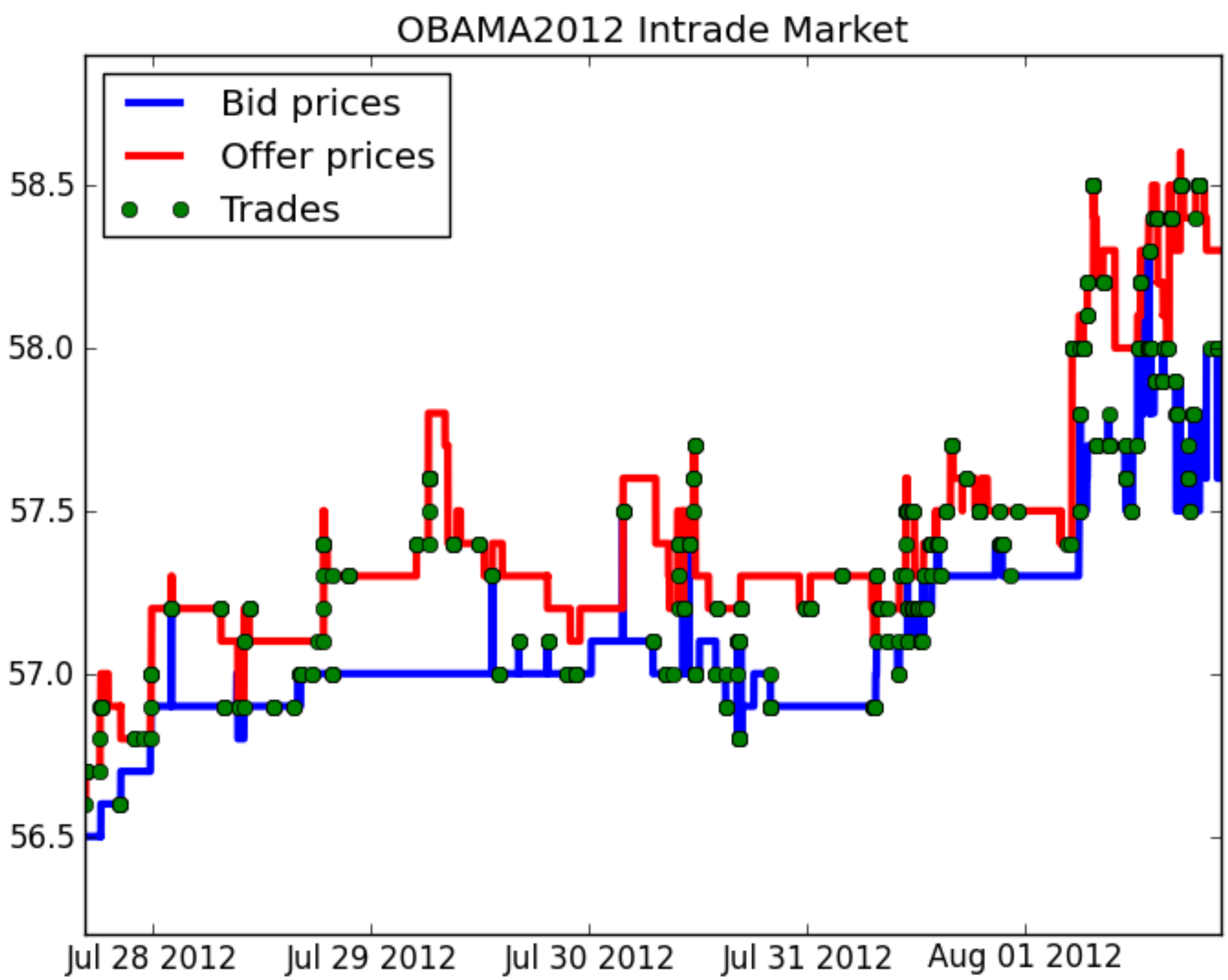
[1] Market Mechanisms: Continuous Double Auction (CDA)

- Used by Intrade.com and Betfair.com
- Market receives a sequence of **orders**
- **Two** types of orders:
  - **Limit** order: trader posts shares to order book
  - **Market** order: trader buys shares in order book

Predict | View All Un-Matched Predictions | Info | Rules

Best (highest) price members are buying at		Best (lowest) price members are selling at	
Price per share	Quantity	Price per share	Quantity
$5.80	86 shares	$5.90	10 shares
$5.78	10 shares	$5.91	22 shares
$5.77	100 shares	$5.92	22 shares
$5.75	9 shares	$5.93	23 shares
$5.74	125 shares	$5.94	22 shares
$5.73	100 shares	$5.95	123 shares
$5.72	6 shares	$5.96	22 shares
$5.71	101 shares	$5.98	8 shares
$5.70	40 shares	$5.99	24 shares
$5.69	10 shares	$6.00	204 shares
$5.68	100 shares	$6.01	23 shares
$5.67	160 shares	$6.02	102 shares
$5.66	22 shares	$6.03	648 shares
$5.65	101 shares	$6.04	122 shares
$5.62	12 shares	$6.05	203 shares

Obama2012 Intrade: Bid+Ask+Trades



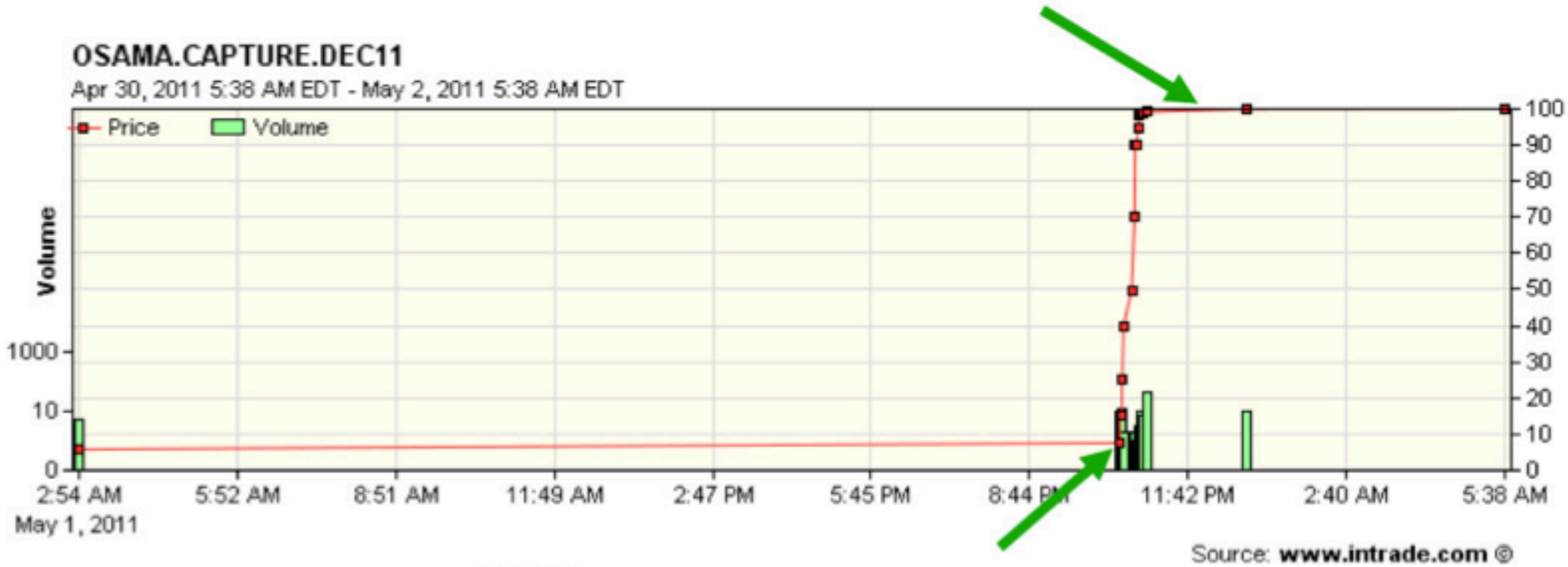
Aside: Problems with the CDA

- Chicken and egg problem: who is willing to join a market if there are no other participants?
- Not a lot of "liquidty": it's very easy to swing prices
- Large bid/ask spreads

- Alternative mechanism: the **automated market maker**, which we will be discussing throughout the 2^{nd} half of the tutorial

[2] How Quickly do Markets Respond?

Obama Addresses the Nation



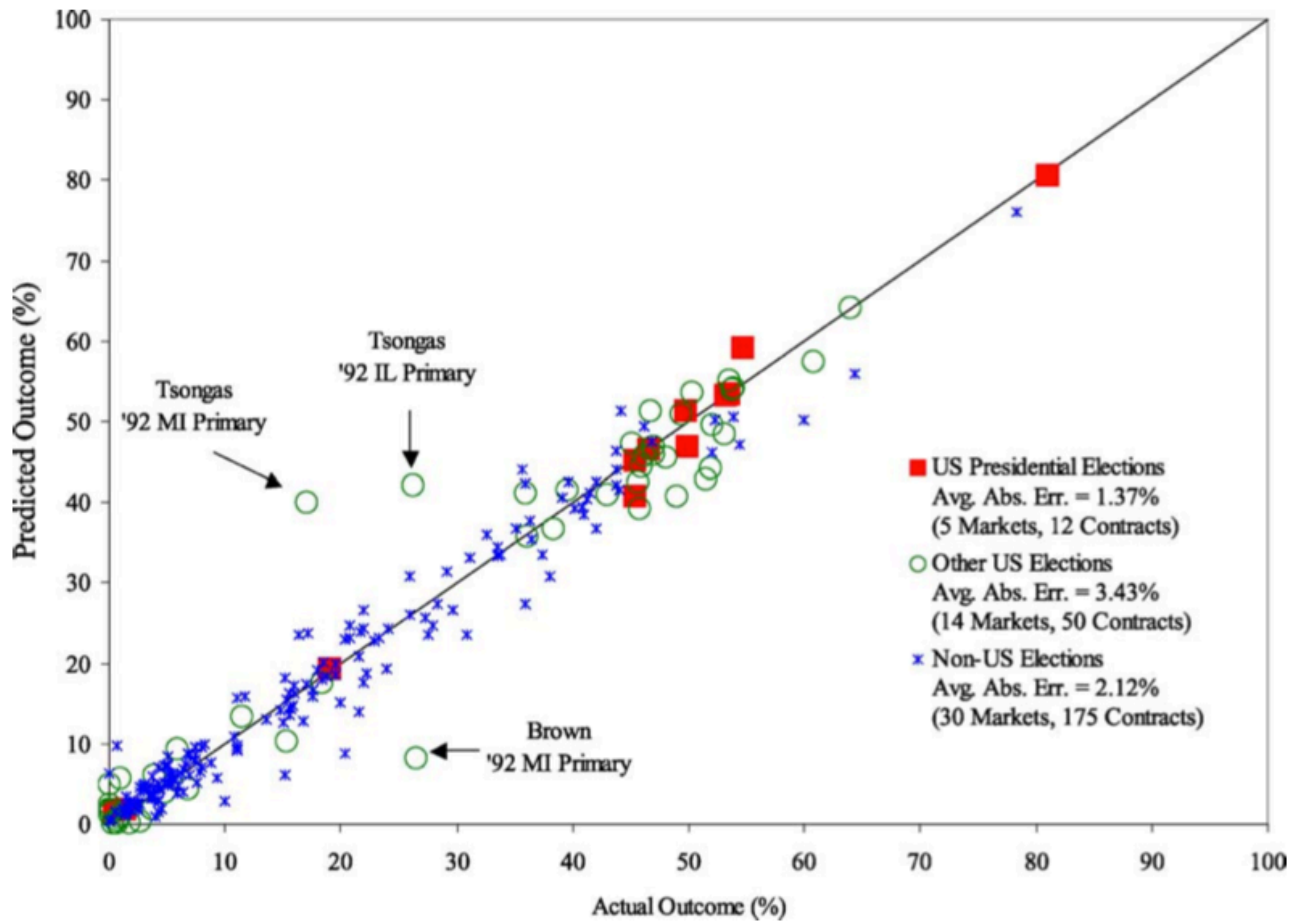
(Donald Rumsfeld's former Chief of Staff)

Source: Snowberg, Wolfers, Zitzewitz 2012

[3] Are Market Prices Accurate?

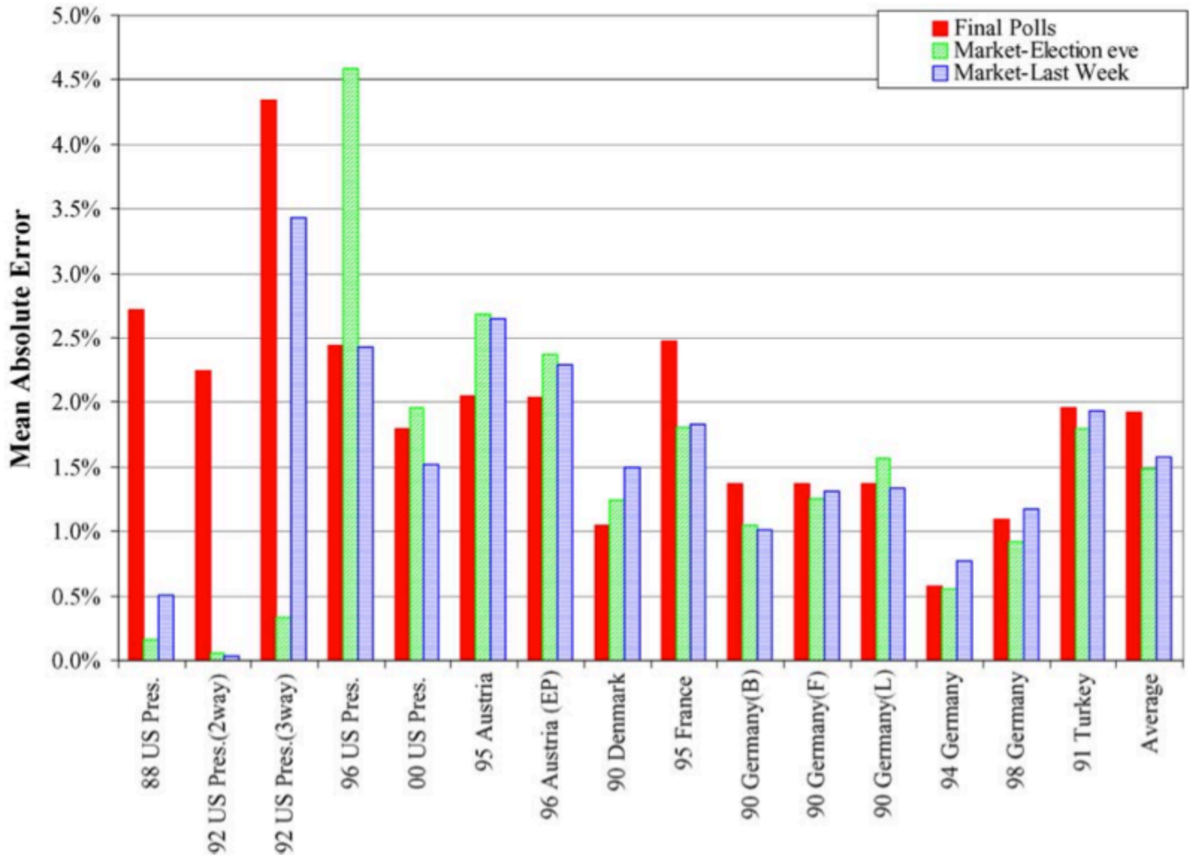
- The market price for Arrow-Debreu security is essentially a “consensus estimate” of the probability of an event
- Are these estimates accurate?
- We can check this on historical data, but…
  - Prices are changing, which price do we use?
  - What is the right metric to measure accuracy?
  - What are we comparing against?

Market Prediction vs. True Vote Share



Berg et al., 2008: "Results From a Dozen Years of Election Futures Markets Research"

Average Polls vs. Market Prices

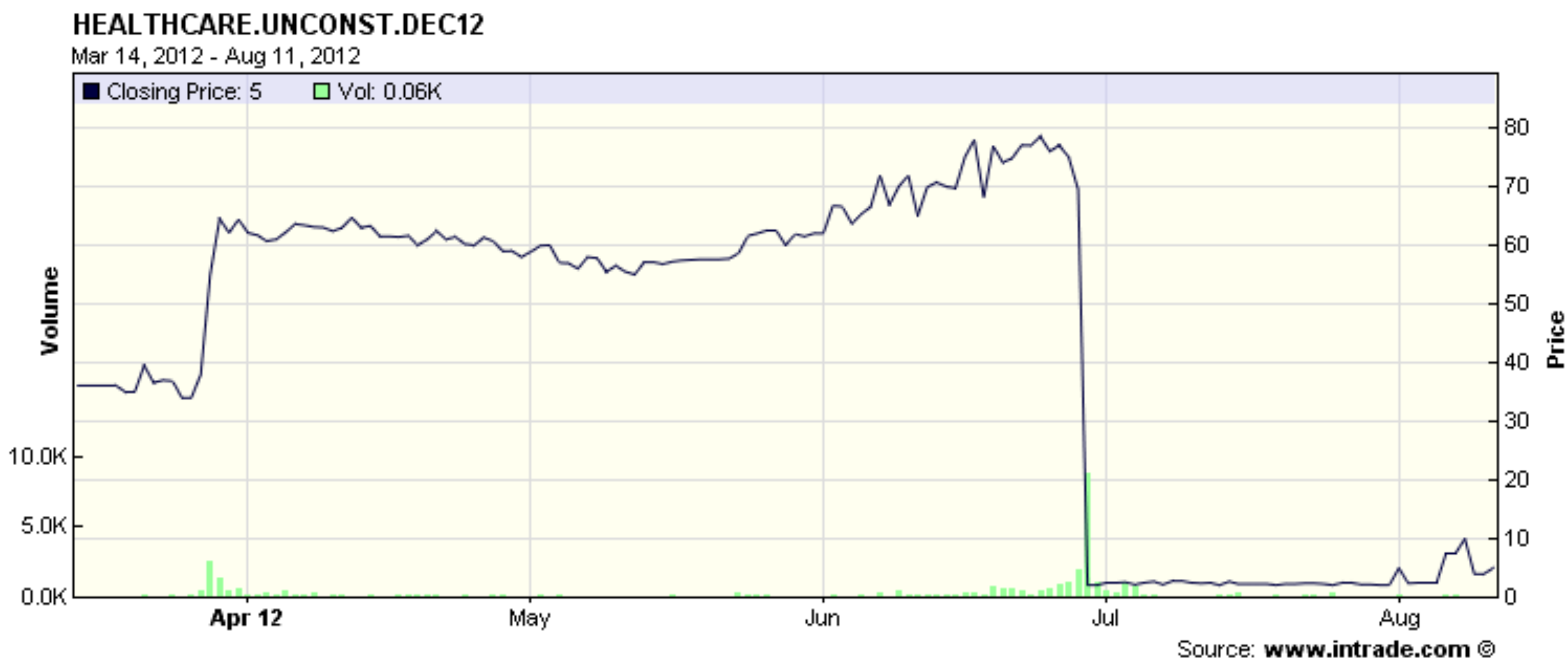


Poll Error: (average from last week)	1.91%
Market Error: (election eve)	1.49%
Market Error: (average from last week)	1.58%

Berg et al., 2008: "Results From a Dozen Years of Election Futures Markets Research"

Aside: Supreme Court + Health Care

Intrade market: "The US Supreme Court to rule individual mandate unconstitutional before midnight ET 31 Dec 2012"



More on Obamacare Prediction Market

- David Leonhardt in the NYTimes: “After several years in which the market was often celebrated as a crystal ball, the Supreme Court ruling was a useful corrective. The prediction-market revolution, like so many others, initially promised more than it could deliver.”
- Response by Robin Hanson on overcomingbias.com: “But the Intrade market on the Obamacare court case was an active valid market, on an appropriate subject. When it assigned a 75% chance to an event it was saying real loud that it would be wrong 1/4 of the time. And studies have consistently found such markets are well-calibrated in this way. What more do you want?”

The Basics: Proper Scoring Rules

1950: Brier on Weather Forecasting

MONTHLY WEATHER REVIEW

EDITOR, JAMES E. CASKEY, JR.

Volume 78
Number 1

JANUARY 1950

Closed March 5, 1950
Issued April 15, 1950

VERIFICATION OF FORECASTS EXPRESSED IN TERMS OF PROBABILITY

GLENN W. BRIER

U. S. Weather Bureau, Washington, D. C.

[Manuscript received February 10, 1950]

INTRODUCTION

Verification of weather forecasts has been a controversial subject for more than a half century. There are a number of reasons why this problem has been so perplexing to meteorologists and others but one of the most important difficulties seems to be in reaching an agreement on the specification of a scale of goodness for weather forecasts. Numerous systems have been proposed but one of the greatest arguments raised against forecast verification is that forecasts which may be the "best" according to the accepted system of arbitrary scores may not be the most

numerically have been discussed previously [1, 2, 3, 4] so that the purpose here will not be to emphasize the enhanced usefulness of such forecasts but rather to point out how some aspects of the verification problem are simplified or solved.

VERIFICATION FORMULA

Suppose that on each of n occasions an event can occur in only one of r possible classes or categories and on one such occasion, i, the forecast probabilities are f_{i1}, f_{i2}, . . . f_{ir}, that the event will occur in classes 1, 2, . . . r, respectively. The r classes are chosen to be mutually

INTRODUCTION

Verification of weather forecasts has been a controversial subject for more than a half century. There are a number of reasons why this problem has been so perplexing to meteorologists and others but one of the most important difficulties seems to be in reaching an agreement on the specification of a scale of goodness for weather forecasts. Numerous systems have been proposed but one of the greatest arguments raised against forecast verification is that forecasts which may be the "best" according to the accepted system of arbitrary scores may not be the most useful forecasts.

TABLE 2.—*Verification of a series of 85 forecasts expressed in terms of the probability of rain*

Forecast probability of rain	Observed proportion of rain cases	Forecast probability of rain	Observed proportion of rain cases
0.00–0.19	0.07	0.60–0.79	0.40
0.20–0.39	.10	0.80–1.00	.50
0.40–0.59	.29		

How Should We Pay a Forecaster?

What is the "right" payment scheme to reward/punish a forecaster who makes a sequence of probability predictions for events that we observe?

- The sequence of outcomes: $y_1, y_2, y_3, \ldots \in \{0,1\}$
- The sequence of forecasts: $p_1, p_2, p_3, \ldots \in [0,1]$
- The forecaster's payment: $\frac{1}{T}\sum_{t=1}^{T} S(y_t, p_t)$

Brier Score ⇔ Quadratic Scoring Rule

- For a binary outcome $y \in \{0,1\}$, $p \in [0,1]$

$$S(y,p) = -(y-p)^2$$

- For one of n outcomes, $y \in \{1,\ldots,n\}$, $\mathbf{p} \in \Delta_n$

$$S(y,\mathbf{p}) = -\sum_{i=1}^{n} (\mathbf{1}_{y=i} - p_i)^2$$

What's Special About This Function?

$$S(y,p) = -(y-p)^2$$

Assume y is random and $\Pr(y = 1) = q$. Then...

$$\underset{p\in[0,1]}{\operatorname{argmax}}\left(\mathrm{E}\left[-(y-p)^2\right]\right)$$

$$= \underset{p\in[0,1]}{\operatorname{argmax}}\left(-q(1-p)^2-(1-q)p^2\right)$$

$$= \underset{p\in[0,1]}{\operatorname{argmax}}\left(-(p-q)^2-q+q^2\right) = q$$

Proper Scoring Rules

- What we have just introduced is the notion of a proper scoring rule, any function S satisfying

$$\mathrm{E}_{y\sim\mathbf{q}}\left[S(y,\mathbf{q})\right] \geq \mathrm{E}_{y\sim\mathbf{q}}\left[S(y,\mathbf{p})\right] \quad \forall \mathbf{p},\mathbf{q} \in \Delta_n$$

- The scoring rule is said to be strictly proper if the above inequality is strict unless $\mathbf{p} = \mathbf{q}$

Another Strictly Proper Scoring Rule

$$S(y,\mathbf{p}) = \log p(y)$$

- This is known as the logarithmic scoring rule. For binary random variables, it can be written as:

$$S(y,p) = \begin{cases} \log p & y = 1 \\ \log(1-p) & y = 0 \end{cases}$$

- **EXERCISE:** check that this is proper!

Scoring Rules == –Loss Functions?

- Effectively, a scoring rule is just a type of loss function
- Scoring rules measure the performance (not loss) of a predicted distribution given a final outcome
- Research on scoring rules is focused more heavily on the incentives of the associated payment mechanism

Designing Scoring Rules
Using Bregman Divergences

Savage 1973

© Journal of the American Statistical Association
December 1971, Volume 66, Number 336
Theory and Methods Section

Elicitation of Personal Probabilities and Expectations

LEONARD J. SAVAGE*

Proper scoring rules, i.e., devices of a certain class for eliciting a person's probabilities and other expectations, are studied, mainly theoretically but with some speculations about application. The relation of proper scoring rules to other economic devices and to the foundations of the personalistic theory of probability is brought out. The implications of various restrictions, especially symmetry restrictions, on scoring rules is explored, usually with a minimum of regularity hypothesis.

1. INTRODUCTION

1.1 Preface

This article is about a class of devices by means of which an idealized *homo economicus*—and therefore, with some approximation, a real person—can be induced to reveal his opinions as expressed by the probabilities that he associates with events or, more generally, his personal expectations of random quantities. My emphasis here is theoretical, though some experimental considerations will be mentioned. The empirical importance of such studies in many areas is now recognized. It was emphasized for the area of economics in an address by Trygve Haavelmo [28, p. 357]:

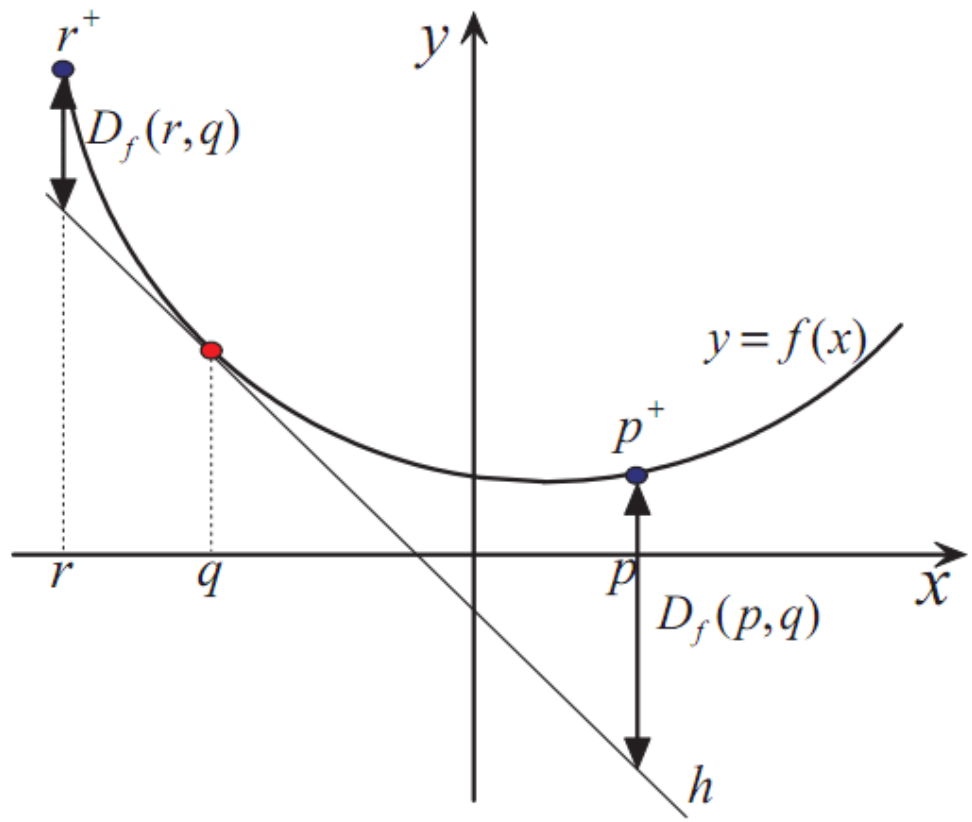
pertaining to it has grown up, some of which will be cited in context and most of which can be found through the references cited, especially the recent and extensive [52] and others that I call "key references."

Bruno de Finetti and I began to write the present article in the spring of 1960, not yet aware of our predecessors and contemporaries. The impetus was de Finetti's, for he had brought us to rediscover McCarthy's [37] insight about convex functions. We expected to make short work of our "little note," but it grew rapidly in many directions and became inordinately delayed. Now we find that the material in the present article is largely mine and that de Finetti has published on diverse aspects of the same subject elsewhere [12, 13, 14, 17]. De Finetti has therefore withdrawn himself from our joint authorship and encouraged me to publish this article alone, though it owes so much to him at every stage, including the final draft.

The article is written for a diverse audience. Consequently, some will find parts of it mathematically too

Digression: Bregman Divergences

- A Bregman divergence measures distance with respect to a convex function f



$$D_f(p,q) = f(p) - f(q) - \nabla f(q) \cdot (p - q)$$

Digression: Bregman Divergences

$$D_f(p,q) = f(p) - f(q) - \nabla f(q)\cdot(p-q)$$

- **Properties:**

$$D_f(p,p) = 0 \quad \forall p$$

$$D_f(p,q) \geq 0 \quad \forall p,q$$

$$D_f(p,q) \neq D_f(q,p) \quad (in\ \ general)$$

Bregman Divergences III

$$D_f(p,q) = f(p) - f(q) - \nabla f(q) \cdot (p - q)$$

- **Example 1**, quadratic:

$$f(p) = \| p \|^2 \Rightarrow D_f(p,q) = \| p - q \|^2$$

- **Example 2**, entropic:

$$f(\mathbf{p}) = \sum_i p_i \log p_i \Rightarrow D_f(\mathbf{p},\mathbf{q}) = \sum_i p_i \log \frac{p_i}{q_i}$$

Bregman Diverg. ⇔ Scoring Rule??

- Let f be any convex function
- Let $\mathbf{e}_i$ be the ith indicator vector, $\mathbf{e}_i = \langle 0,\ldots,0,1,0,\ldots,0\rangle$
- Let **p, q** be any two distributions
- Fact: There exists a function g such that

$$\mathrm{E}_{i\sim\mathbf{q}}\left[D_f(\mathbf{e}_i,\mathbf{p})\right] = D_f(\mathbf{q},\mathbf{p}) + g(\mathbf{q})$$

and so

$$\operatorname*{arg\,max}_{\mathbf{p}\in\Delta_n} \mathrm{E}_{i\sim\mathbf{q}}\left[-D_f(\mathbf{e}_i,\mathbf{p})\right] = \mathbf{q}$$

- This is the scoring rule property!!

Bregman Diverg. ⇔ Scoring Rule!!

- We now have a recipe for constructing scoring rules: Take any convex function f and set

$$S(i,\mathbf{p}) = -D_f(\mathbf{e}_i,\mathbf{p})$$

- Quadratic Scoring Rule: $f(\mathbf{p}) = \|\mathbf{p}\|_2^2$

- Log Scoring Rule: $f(\mathbf{p}) = \sum_i p_i \log p_i$

Brief Literature Review

Surrogate regret bounds for proper losses
MD Reid, RC Williamson
Proceedings of the 26th Annual International Conference on Machine Learning ...

Information, divergence and risk for binary experiments
MD Reid, RC Williamson
Journal of Machine Learning Research 12, 731-817

Composite binary losses
MD Reid, RC Williamson
The Journal of Machine Learning Research 9999, 2387-2422

Composite Multiclass Losses
E Vernet, ENS Cachan, RC Williamson, MD Reid
Neural Information Processing Systems

Mixability is Bayes Risk Curvature Relative to Log Loss
T van Erven, MD Reid, RC Williamson
Proceedings of the 24th Annual Conference on Learning Theory (COLT)

Market Scoring Rules
for Belief Aggregation

Learning a Consensus?

- Scoring rules are useful for incentivizing one individual to state his beliefs about a probability, but what if we'd like to learn from a crowd

- **Proposal**: We could just pay every individual according to a scoring rule.
- **Problems**:
  - This could be very expensive!
  - How should we combine estimates?
  - How can we weed out noise traders?
  - How can we weed out copycats?

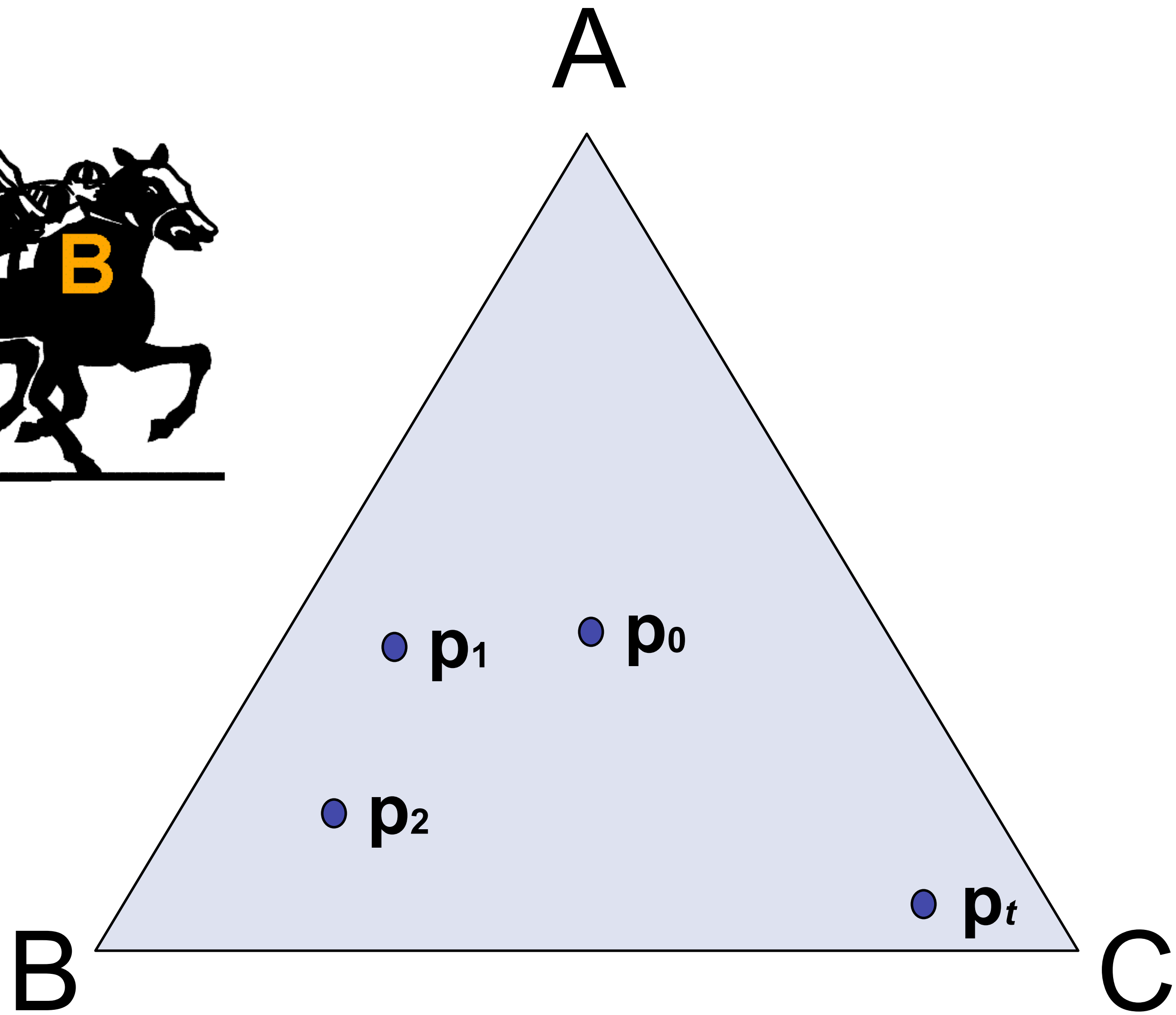
Market Scoring Rules

Robin Hanson proposed the following idea to create a prediction market based on an automated market maker:

- Suppose we have a random variable X which will take one of n values $\{1, 2, \ldots, n\}$
- The MM chooses a scoring rule S and announces it
- The MM then posts an initial distribution (prior) $\mathbf{p}_0$
- Traders arrive, one-by-one, giving updates $\mathbf{p}_{t-1} \rightarrow \mathbf{p}_t$
- Eventually, outcome X is revealed, and trader t earns (or loses)

$$S(X, \mathbf{p}_t) - S(X, \mathbf{p}_{t-1})$$

Market Scoring Rule



Incentives and Costs

- Assume trader t has belief distribution $\mathbf{p}$ on X, which can (and should!) depend on previous market observations
- Suppose he wants to maximize his payment

$$\operatorname{argmax}_{\mathbf{p}_t} \mathrm{E}_{X\sim\mathbf{p}}[S(X,\mathbf{p}_t) - S(X,\mathbf{p}_{t-1})]$$

always non-negative!

$$= \operatorname{argmax}_{\mathbf{p}_t} \mathrm{E}_{X\sim\mathbf{p}}[S(X,\mathbf{p}_t)] = \mathbf{p}$$

- The MM must make all payments, which total

$$\sum_{t=1}^{T}\left[S(X,\mathbf{p}_t) - S(X,\mathbf{p}_{t-1})\right] = S(X,\mathbf{p}_T) - S(X,\mathbf{p}_0)$$

- This is bounded! This is like MM's subsidy to market.

LMSR: Log Market Scoring Rule

- Initial hypothesis $\mathbf{p}_0$ is the uniform distribution
- Trader t posts an update $\mathbf{p}_{t-1} \rightarrow \mathbf{p}_t$
- After X is revealed, trader t earns $\log(\mathbf{p}_t(X)/\mathbf{p}_{t-1}(X))$

- Hanson: the LMSR is an important special case, the only MSR for which "betting on conditional probabilities does not affect marginal probabilities"

- The market maker's worst case loss is bounded by $\log n$, where n is the number of possible values of X

Prediction, Belief, and Markets: Part 2

http://aaaimarketstutorial.pbworks.com

Jake Abernethy, UPenn → UMich
Jenn Wortman Vaughan, Microsoft Research NYC

Outline of Part 2

1. Automated market makers for securities markets
2. Designing markets for large outcome spaces
3. Duality & connections to online learning
4. Recent extensions and additional topics

Back to Securities Markets

Arrow–Debreu Securities

Barack Obama to be re-elected President in 2012

Last prediction was: **$5.56** / share

Today's Change: ▲ **+$0.06** (+1.1%)

Contract Type: **0-100**

55.6%

CHANCE

Arrow–Debreu Securities

Barack Obama to be re-elected President in 2012

Last prediction was **$5.56** / share → **55.6%** CHANCE

Today's Change: ▲ **+$0.06** (+1.1%)

Contract Type: **0-100**

Potential payoff is \$10. If I think that the probability of this event is p, I should

- Buy this security at any price less than \$10p
- Sell this security at any price greater than \$10p

Current price measures the population's collective beliefs

Arrow–Debreu Securities

Barack Obama to be re-elected President in 2012

Last prediction was **$5.56** / share → **55.6%** CHANCE

Today's Change: ▲ **+$0.06** (+1.1%)

Contract Type: **0-100**

How do we arrive at the current price?

Arrow–Debreu Securities

How do we arrive at the current price?

- Traditional stock market style pricing (continuous double auction) – low liquidity, huge spreads

Arrow–Debreu Securities

How do we arrive at the current price?

- Traditional stock market style pricing (continuous double auction) – low liquidity, huge spreads
- Automated market maker – willing to risk a (bounded) loss in order to encourage trades

Market Makers for Complete Markets

- In a complete market, a security is offered for each of a set of mutually exclusive and exhaustive events

Market Makers for Complete Markets

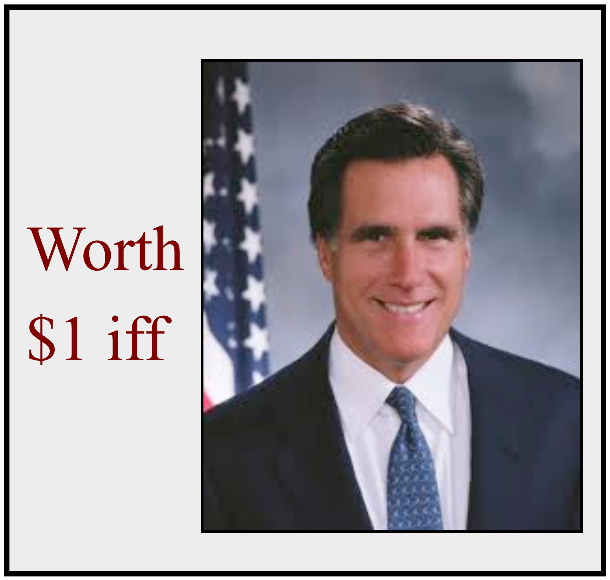
- In a complete market, a security is offered for each of a set of mutually exclusive and exhaustive events

Worth $1 iff

Worth $1 iff

Market Makers for Complete Markets

- In a complete market, a security is offered for each of a set of mutually exclusive and exhaustive events

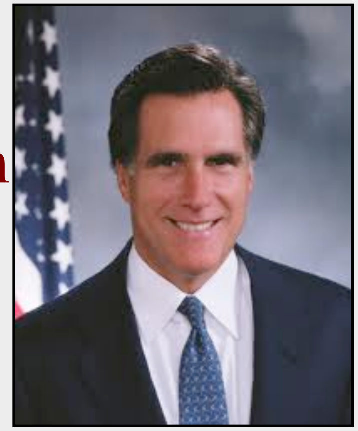


- An automated market maker is always willing to buy and sell these securities at some price

Cost Functions

Cost Functions

Worth $1 iff

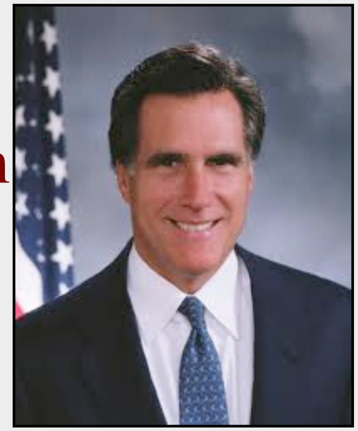


Worth $1 iff

Already purchased: q_1 shares q_2 shares

Cost Functions

Worth $1 iff



Worth $1 iff

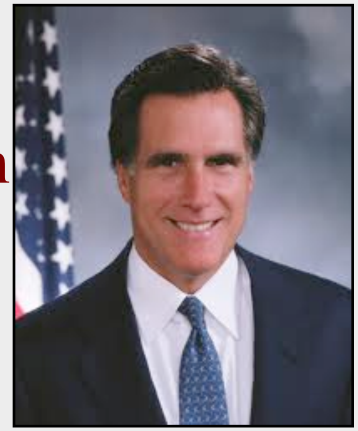
Already purchased:	q_1 shares	q_2 shares
Want to purchase:	r_1 shares	r_2 shares

Cost Functions

Cost of purchase:

$$C(\mathbf{q} + \mathbf{r}) - C(\mathbf{q})$$

Worth $1 iff



Worth $1 iff

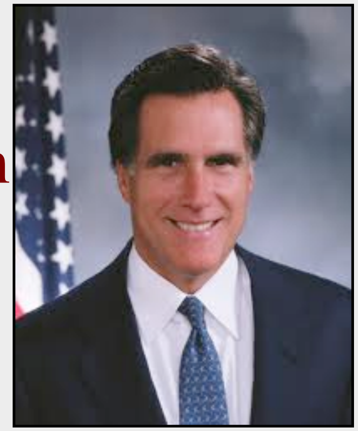
Already purchased:	q_1 shares	q_2 shares
Want to purchase:	r_1 shares	r_2 shares

Cost Functions

Cost of purchase:

$$C(\mathbf{q} + \mathbf{r}) - C(\mathbf{q})$$

Worth $1 iff



Worth $1 iff

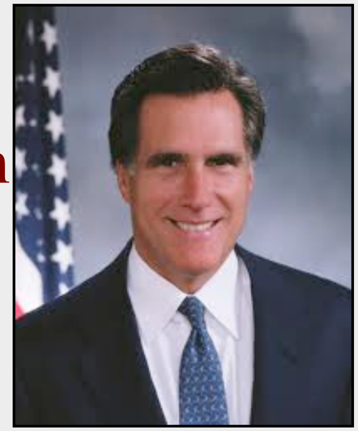
Already purchased:	q_1 shares	q_2 shares
Want to purchase:	r_1 shares	r_2 shares
Instantaneous prices:	$p_1 = \partial C / \partial q_1$	$p_2 = \partial C / \partial q_2$

Cost Functions

Cost of purchase:

$$C(\mathbf{q} + \mathbf{r}) - C(\mathbf{q})$$

Worth $1 iff



Worth $1 iff

Already purchased:	q_1 shares	q_2 shares
Want to purchase:	r_1 shares	r_2 shares
Instantaneous prices:	$p_1 = \partial C / \partial q_1$	$p_2 = \partial C / \partial q_2$

"predictions"

Back to the LMSR

Remember the logarithmic market scoring rule…

- Initial hypothesis $\mathbf{p}_0$ is the uniform distribution

- Trader t posts an update $\mathbf{p}_{t\text{-}1} \rightarrow \mathbf{p}_t$

- After outcome i is revealed, trader t receives

$$\log(p_{t,i}) - \log(p_{t\text{-}1,i}) = \log(p_{t,i} / p_{t\text{-}1,i})$$

Back to the LMSR

The logarithmic market scoring rule can be implemented as a cost function based market with cost function

$$C(q_1,\ldots,q_N) = \log \sum_{i=1}^{N} \exp(q_i)$$

and instantaneous prices

$$p_i = \frac{\exp(q_i)}{\Sigma_j \exp(q_j)}$$

Back to the LMSR

The logarithmic market scoring rule can be implemented as a cost function based market with cost function

$$C(q_1,\ldots,q_N) = \log \sum_{i=1}^{N} \exp(q_i)$$

and instantaneous prices

$$p_i = \frac{\exp(q_i)}{\Sigma_j \exp(q_j)}$$

Notice that p_i is increasing in q_i and the prices sum to 1

Equivalence

For all $\mathbf{p}$, $\mathbf{p'}$, $\mathbf{q}$, $\mathbf{q'}$, such that $\nabla C(\mathbf{q}) = \mathbf{p}$ and $\nabla C(\mathbf{q'}) = \mathbf{p'}$, for all outcomes i, a trader who changed the market state from $\mathbf{p}$ to $\mathbf{p'}$ in the MSR would receive the same total payoff as a trader who changed the market state from $\mathbf{q}$ to $\mathbf{q'}$ in the cost function based market.

[Hanson 03; Chen & Pennock 07]

A Proof in One Slide

A Proof in One Slide

cost function payoff

$$= \left(q'_i - q_i\right) - \left(C(\mathbf{q}') - C(\mathbf{q})\right)$$

security value when the outcome *i* occurs

cost of the purchase

A Proof in One Slide

cost function payoff

$$= \left(q'_i - q_i\right) - \left(C(\mathbf{q}') - C(\mathbf{q})\right)$$

A Proof in One Slide

cost function payoff

$$= \left(q'_i - q_i\right) - \left(C(\mathbf{q}') - C(\mathbf{q})\right)$$

$$= \left(q'_i - q_i\right) - \left(\log \sum_j e^{q'_j} - \log \sum_j e^{q_j}\right)$$

by definition

A Proof in One Slide

cost function payoff

$$= \left(q'_i - q_i\right) - \left(C(\mathbf{q}') - C(\mathbf{q})\right)$$

$$= \left(q'_i - q_i\right) - \left(\log \sum_j e^{q'_j} - \log \sum_j e^{q_j}\right)$$

A Proof in One Slide

cost function payoff

$$= \left(q'_i - q_i\right) - \left(C(\mathbf{q}') - C(\mathbf{q})\right)$$

$$= \left(\log e^{q'_i} - \log e^{q_i}\right) - \left(\log \sum_j e^{q'_j} - \log \sum_j e^{q_j}\right)$$

A Proof in One Slide

cost function payoff

$$= \left(q'_i - q_i\right) - \left(C(\mathbf{q}') - C(\mathbf{q})\right)$$

$$= \left(\log e^{q'_i} - \log e^{q_i}\right) - \left(\log \sum_j e^{q'_j} - \log \sum_j e^{q_j}\right)$$

$$= \log \frac{e^{q'_i}}{\sum_j e^{q'_j}} - \log \frac{e^{q_i}}{\sum_j e^{q_j}}$$

A Proof in One Slide

cost function payoff

$$= (q'_i - q_i) - (C(\mathbf{q}') - C(\mathbf{q}))$$

$$= \left(\log e^{q'_i} - \log e^{q_i}\right) - \left(\log \sum_j e^{q'_j} - \log \sum_j e^{q_j}\right)$$

$$= \log \frac{e^{q'_i}}{\sum_j e^{q'_j}} - \log \frac{e^{q_i}}{\sum_j e^{q_j}}$$

prices!

A Proof in One Slide

cost function payoff

$$= \left(q'_i - q_i\right) - \left(C(\mathbf{q'}) - C(\mathbf{q})\right)$$

$$= \left(\log e^{q'_i} - \log e^{q_i}\right) - \left(\log \sum_j e^{q'_j} - \log \sum_j e^{q_j}\right)$$

$$= \log \frac{e^{q'_i}}{\sum_j e^{q'_j}} - \log \frac{e^{q_i}}{\sum_j e^{q_j}}$$

$$= \log p'_i - \log p_i$$

A Proof in One Slide

cost function payoff

$$= \left(q'_i - q_i\right) - \left(C(\mathbf{q}') - C(\mathbf{q})\right)$$

$$= \left(\log e^{q'_i} - \log e^{q_i}\right) - \left(\log \sum_j e^{q'_j} - \log \sum_j e^{q_j}\right)$$

$$= \log \frac{e^{q'_i}}{\sum_j e^{q'_j}} - \log \frac{e^{q_i}}{\sum_j e^{q_j}}$$

$$= \log p'_i - \log p_i$$

= scoring rule payoff

More Generally

- Any market scoring rule can be implemented as a cost function based market [Chen & Pennock 07; Chen & Vaughan 10; Abernethy & Frongillo 11; …]

More Generally

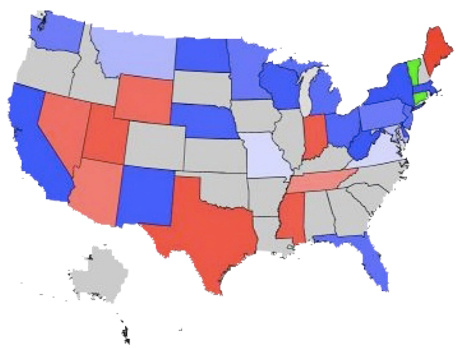
- Any market scoring rule can be implemented as a cost function based market [Chen & Pennock 07; Chen & Vaughan 10; Abernethy & Frongillo 11; …]

- Advantages:
  - Retains the good incentive properties of the MSR
  - Arguably more “natural” for traders
  - Exposure to risk is more transparent

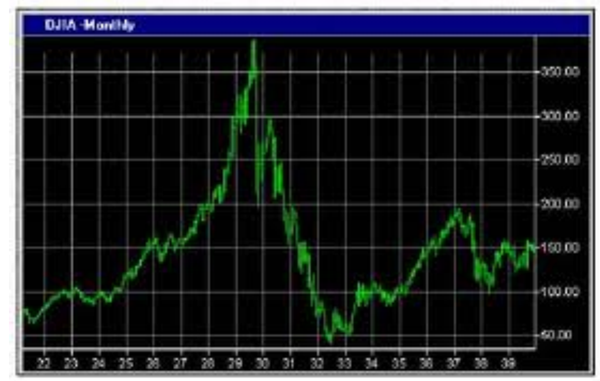
Beyond Complete Markets

Complex Outcome Spaces

$n!$



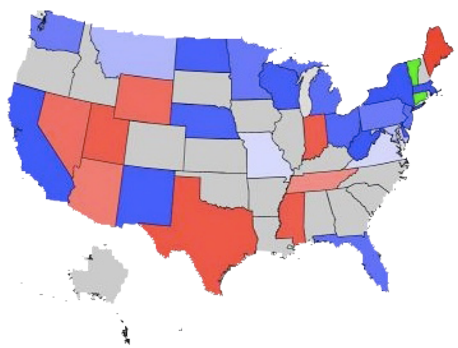
2^n



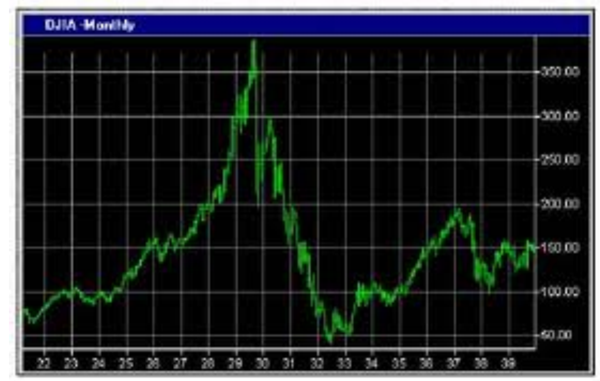
Infinite

Complex Outcome Spaces

$n!$



2^n

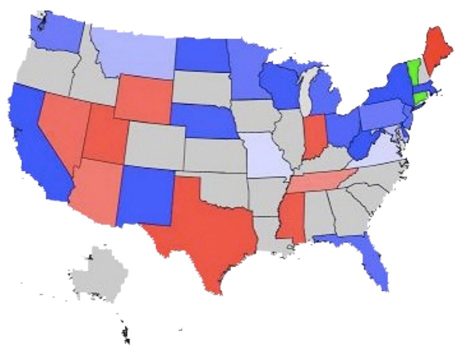


Infinite

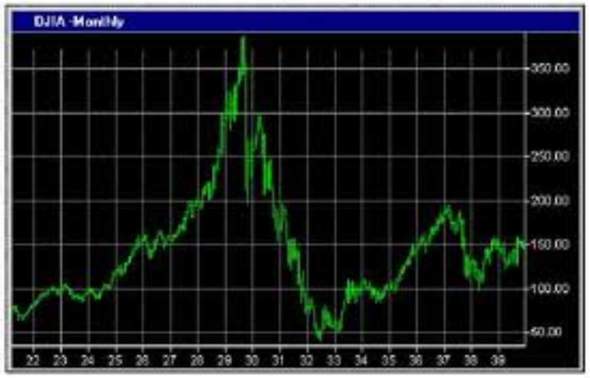
MSR-NYC's WiseQ
[Dudik et al., 2013]

Complex Outcome Spaces

$n!$



2^n

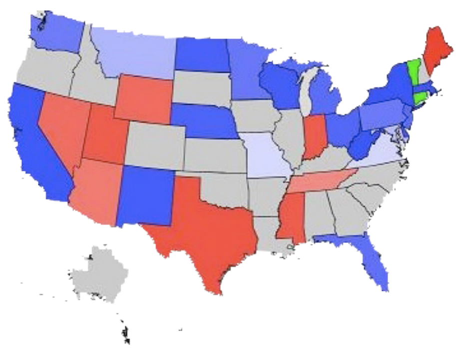


Infinite

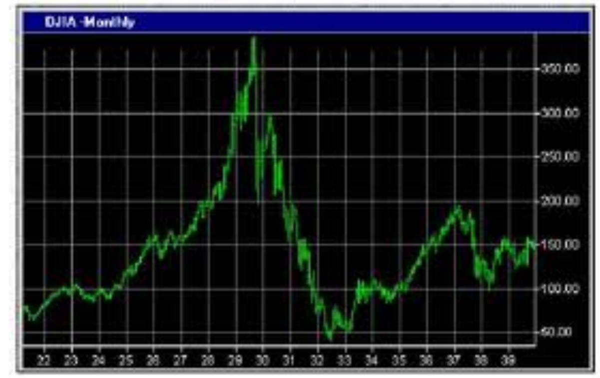
- Cannot simply run a standard market like LMSR
  - Calculating prices is intractable [Chen et al., 2008]
  - Reasoning about probabilities is too hard for traders

Complex Outcome Spaces

$n!$



2^n



Infinite

- Cannot simply run a standard market like LMSR
  - Calculating prices is intractable [Chen et al., 2008]
  - Reasoning about probabilities is too hard for traders
- Can run separate, independent markets (e.g., horses to win, place, or show) but this ignores logical dependences

Complex Outcome Spaces

Given a small set of securities over a very large (or infinite) state space, how can we design a consistent market that can be operated efficiently?

[Abernethy, Chen, and Vaughan, EC 2011;
long version in ACM TEAC 2013]

Menu of Securities

We would like to offer a menu of securities $\{1, \ldots, K\}$ specified by a payoff function ρ

Menu of Securities

We would like to offer a menu of securities $\{1, \ldots, K\}$ specified by a payoff function ρ

payoff

securities

outcomes

10	0	5.5	0	17	0
.9	.9	.9	0	.9	.9
0	42	0	10	10	10
0	0	11.5	8	0	0
1	0	0	0	0	1

Example: Pair Betting

\$1 if and only if horse i finishes ahead of horse j

Example: Pair Betting

\$1 if and only if horse *i* finishes ahead of horse *j*

	A<B	B<A	A<C	C<A	B<C	C<B
ABC	1	0	1	0	1	0
ACB	1	0	1	0	0	1
BAC	0	1	1	0	1	0
BCA	0	1	0	1	1	0
CAB	1	0	0	1	0	1
CBA	0	1	0	1	0	1

What are "reasonable" prices?

What are “reasonable” prices?

For complete markets…

What are "reasonable" prices?

For complete markets… $$\sum_i p_i = 1$$

What are "reasonable" prices?

For complete markets… $$\sum_i p_i = 1$$

For pair betting…

What are "reasonable" prices?

For complete markets…

$$\sum_i p_i = 1$$

For pair betting…

$$p_{i<j} + p_{j<i} = 1$$

What are "reasonable" prices?

For complete markets… $$\sum_i p_i = 1$$

For pair betting… $$p_{i<j} + p_{j<i} = 1$$

$$1 \le p_{i<j} + p_{j<k} + p_{k<i} \le 2$$

What are "reasonable" prices?

For complete markets… $$\sum_i p_i = 1$$

For pair betting… $$p_{i<j} + p_{j<i} = 1$$

$$1 \le p_{i<j} + p_{j<k} + p_{k<i} \le 2$$

what else?

What are "reasonable" prices?

For complete markets… $\sum_i p_i = 1$

For pair betting… $p_{i<j} + p_{j<i} = 1$

$$1 \le p_{i<j} + p_{j<k} + p_{k<i} \le 2$$

what else?

In general… ???

An Axiomatic Approach

Path independence: The cost of acquiring a bundle **r** of securities must be the same no matter how the trader splits up the purchase.

An Axiomatic Approach

Path independence: The cost of acquiring a bundle **r** of securities must be the same no matter how the trader splits up the purchase. Formally,

$$\text{Cost}(\mathbf{r} + \mathbf{r}' \mid \mathbf{r_1}, \mathbf{r_2}, \ldots, \mathbf{r_t})$$
$$= \text{Cost}(\mathbf{r} \mid \mathbf{r_1}, \mathbf{r_2}, \ldots, \mathbf{r_t}) + \text{Cost}(\mathbf{r}' \mid \mathbf{r_1}, \mathbf{r_2}, \ldots, \mathbf{r_t}, \mathbf{r})$$

An Axiomatic Approach

Path independence: The cost of acquiring a bundle $\mathbf{r}$ of securities must be the same no matter how the trader splits up the purchase. Formally,

$$\text{Cost}(\mathbf{r} + \mathbf{r'} \mid \mathbf{r_1}, \mathbf{r_2}, \ldots, \mathbf{r_t})$$
$$= \text{Cost}(\mathbf{r} \mid \mathbf{r_1}, \mathbf{r_2}, \ldots, \mathbf{r_t}) + \text{Cost}(\mathbf{r'} \mid \mathbf{r_1}, \mathbf{r_2}, \ldots, \mathbf{r_t}, \mathbf{r})$$

This alone implies the existence of a cost potential function!

$$\text{Cost}(\mathbf{r} \mid \mathbf{r_1}, \mathbf{r_2}, \ldots, \mathbf{r_t})$$
$$= C(\mathbf{r_1} + \mathbf{r_2} + \ldots + \mathbf{r_t} + \mathbf{r}) - C(\mathbf{r_1} + \mathbf{r_2} + \ldots + \mathbf{r_t})$$

An Axiomatic Approach

- **Existence of instantaneous prices:** *C* must be continuous and differentiable

An Axiomatic Approach

- **Existence of instantaneous prices:** C must be continuous and differentiable

- **Information incorporation:** The purchase of a bundle **r** should never cause the price of **r** to decrease

An Axiomatic Approach

- **Existence of instantaneous prices:** C must be continuous and differentiable
- **Information incorporation:** The purchase of a bundle **r** should never cause the price of **r** to decrease
- **No arbitrage:** It is never possible to purchase a bundle **r** with a guaranteed positive profit regardless of outcome

An Axiomatic Approach

- **Existence of instantaneous prices:** C must be continuous and differentiable
- **Information incorporation:** The purchase of a bundle **r** should never cause the price of **r** to decrease
- **No arbitrage:** It is never possible to purchase a bundle **r** with a guaranteed positive profit regardless of outcome
- **Expressiveness:** A trader must always be able to set the market prices to reflect his beliefs

An Axiomatic Approach

Theorem: Under these five conditions, costs must be determined by a convex cost function C such that

$$\{\nabla C(\mathbf{q}) : \mathbf{q} \in \mathrm{R}^K\} \approx \mathrm{Hull}(\boldsymbol{\rho})$$

An Axiomatic Approach

Theorem: Under these five conditions, costs must be determined by a convex cost function C such that

$$\{\nabla C(\mathbf{q}) : \mathbf{q} \in \mathbb{R}^K\} \approx \text{Hull}(\boldsymbol{\rho})$$

reachable price vectors

An Axiomatic Approach

Theorem: Under these five conditions, costs must be determined by a convex cost function C such that

$$\{\nabla C(\mathbf{q}) : \mathbf{q} \in \mathrm{R}^K\} \approx \mathrm{Hull}(\boldsymbol{\rho})$$

reachable price vectors

securities

outcomes

10	0	5.5	0	17	0
.9	.9	.9	0	.9	.9
0	42	0	10	10	10
0	0	11.5	8	0	0
1	0	0	0	0	1

An Axiomatic Approach

Theorem: Under these five conditions, costs must be determined by a convex cost function C such that

$$\{\nabla C(\mathbf{q}) : \mathbf{q} \in \mathrm{R}^K\} \approx \mathrm{Hull}(\boldsymbol{\rho})$$

reachable price vectors

securities

outcomes

1	0	0	0	0
0	1	0	0	0
0	0	1	0	0
0	0	0	1	0
0	0	0	0	1

Cost Functions Via Duality & The Connection to Online Learning

How do we find these cost functions?

- **Fact:** A closed, differentiable function C is convex if and only if it can be written in the form

$$C(\mathbf{q}) = \sup_{\mathbf{x} \in \mathrm{dom}(R)} \mathbf{x} \cdot \mathbf{q} - R(\mathbf{x})$$

for a strictly convex function R called the conjugate.

How do we find these cost functions?

- **Fact:** A closed, differentiable function C is convex if and only if it can be written in the form

$$C(\mathbf{q}) = \sup_{\mathbf{x} \in \mathrm{dom}(R)} \mathbf{x} \cdot \mathbf{q} - R(\mathbf{x})$$

for a strictly convex function R called the conjugate.

Furthermore, $\nabla C(\mathbf{q}) = \arg\max_{\mathbf{x} \in \mathrm{dom}(R)} \mathbf{x} \cdot \mathbf{q} - R(\mathbf{x})$

How do we find these cost functions?

- **Fact:** A closed, differentiable function C is convex if and only if it can be written in the form

$$C(\mathbf{q}) = \sup_{\mathbf{x} \in \mathrm{dom}(R)} \mathbf{x}\cdot\mathbf{q} - R(\mathbf{x})$$

for a strictly convex function R called the conjugate.

Furthermore, $\nabla C(\mathbf{q}) = \arg\max_{\mathbf{x} \in \mathrm{dom}(R)} \mathbf{x}\cdot\mathbf{q} - R(\mathbf{x})$

To generate a convex cost function C, we just have to choose an appropriate conjugate function and domain!

But how do we choose R?

But how do we choose *R*?

We can borrow ideas from online linear optimization (or the simpler expert advice setting) and in particular, Follow the Regularized Leader algorithms

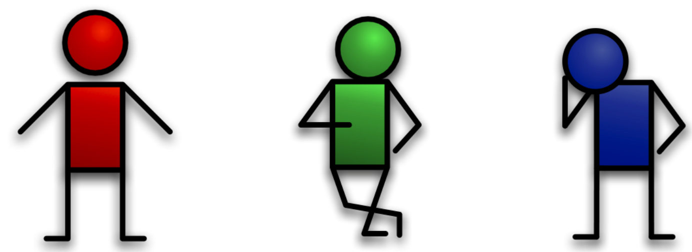
- Market's conjugate function ≈ regularizer

Learning from Expert Advice

Suppose we would like to choose actions based on the advice of n "experts" (people, algorithms, features…)

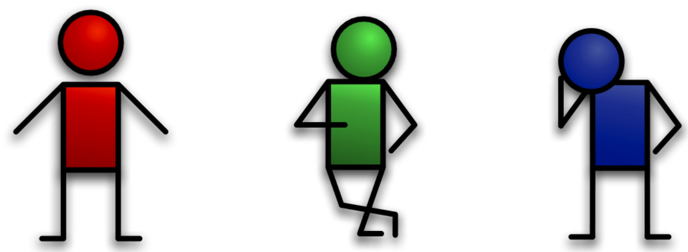
Learning from Expert Advice

Suppose we would like to choose actions based on the advice of n "experts" (people, algorithms, features…)



Learning from Expert Advice

Suppose we would like to choose actions based on the advice of *n* "experts" (people, algorithms, features…)

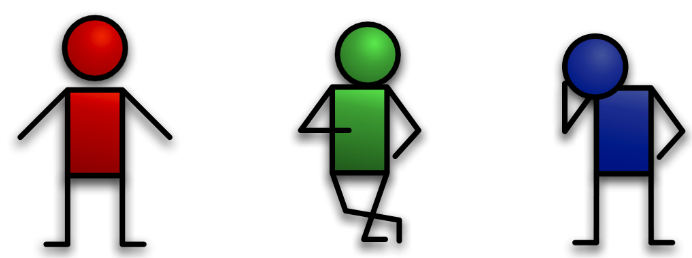


At each round *t*,

- Algorithm selects weights $w_{t,1}$ $w_{t,2}$ $w_{t,3}$

Learning from Expert Advice

Suppose we would like to choose actions based on the advice of n "experts" (people, algorithms, features…)

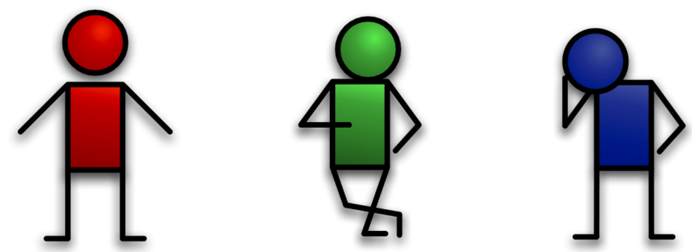


At each round t,

- Algorithm selects weights $w_{t,1}$ $w_{t,2}$ $w_{t,3}$
- Experts suffer a loss $l_{t,1}$ $l_{t,2}$ $l_{t,3}$

Learning from Expert Advice

Suppose we would like to choose actions based on the advice of n "experts" (people, algorithms, features…)



At each round t,

- Algorithm selects weights $w_{t,1}$ $w_{t,2}$ $w_{t,3}$
- Experts suffer a loss $l_{t,1}$ $l_{t,2}$ $l_{t,3}$
- Algorithm suffers a loss $\mathbf{w}_t \cdot \mathbf{l}_t$

What is the goal?

- Ideally, we'd like to bound the cumulative loss $\sum_{t=1}^{T} \mathbf{w}_t \cdot \mathbf{l}_t$

What is the goal?

- Ideally, we'd like to bound the cumulative loss $\sum_{t=1}^{T} \mathbf{w}_t \cdot \mathbf{l}_t$
- Instead, we look at the algorithm's regret

$$\sum_{t=1}^{T} \mathbf{w}_t \cdot \mathbf{l}_t - \min_{\mathbf{w} \in K} \mathbf{w} \cdot \mathbf{L}_T$$

What is the goal?

- Ideally, we'd like to bound the cumulative loss $\sum_{t=1}^{T} \mathbf{w}_t \cdot \mathbf{l}_t$

- Instead, we look at the algorithm's regret

$$\sum_{t=1}^{T} \mathbf{w}_t \cdot \mathbf{l}_t - \min_{\mathbf{w} \in K} \mathbf{w} \cdot \mathbf{L}_T$$

algorithm's loss

What is the goal?

- Ideally, we'd like to bound the cumulative loss $\sum_{t=1}^{T} \mathbf{w}_t \cdot \mathbf{l}_t$
- Instead, we look at the algorithm's regret

$$\sum_{t=1}^{T} \mathbf{w}_t \cdot \mathbf{l}_t - \min_{\mathbf{w} \in K} \mathbf{w} \cdot \mathbf{L}_T$$

cumulative loss vector

algorithm's loss

loss of the best fixed weight vector in hindsight

What is the goal?

- Ideally, we'd like to bound the cumulative loss $\sum_{t=1}^{T} \mathbf{w}_t \cdot \mathbf{l}_t$
- Instead, we look at the algorithm's regret

$$\sum_{t=1}^{T} \mathbf{w}_t \cdot \mathbf{l}_t - \min_{\mathbf{w} \in K} \mathbf{w} \cdot \mathbf{L}_T$$

What is the goal?

- Ideally, we'd like to bound the cumulative loss $\sum_{t=1}^{T} \mathbf{w}_t \cdot \mathbf{l}_t$

- Instead, we look at the algorithm's regret

$$\sum_{t=1}^{T} \mathbf{w}_t \cdot \mathbf{l}_t - \min_{\mathbf{w} \in K} \mathbf{w} \cdot \mathbf{L}_T$$

- Can achieve optimal ($O(T^{1/2})$) regret with Follow the Regularized Leader

$$\mathbf{w}_{t+1} = \arg\min_{\mathbf{w} \in K} \mathbf{w} \cdot \mathbf{L}_t + R(\mathbf{w})$$

cumulative loss regularizer

Online Linear Opt.

Market Making

Online Linear Opt.

- Learner maintains weights $\mathbf{w}_t \in K$ over n items/experts

Market Making

- Market maker maintains prices $\mathbf{p}_t \in \Pi$ over n contracts

Online Linear Opt.

- Learner maintains weights $\mathbf{w}_t \in K$ over n items/experts
- Items have loss vector $\mathbf{l}_t$, cumulatively $\mathbf{L}_{t+1} = \mathbf{L}_t + \mathbf{l}_{t+1}$

Market Making

- Market maker maintains prices $\mathbf{p}_t \in \Pi$ over n contracts
- Contracts are purchased in bundles $\mathbf{r}_t$, and $\mathbf{q}_{t+1} = \mathbf{q}_t + \mathbf{r}_{t+1}$

Online Linear Opt.

- Learner maintains weights $\mathbf{w}_t \in K$ over n items/experts
- Items have loss vector $\mathbf{l}_t$, cumulatively $\mathbf{L}_{t+1} = \mathbf{L}_t + \mathbf{l}_{t+1}$
- FTRL selects weights

$$\mathbf{w}_{t+1} = \arg\min_{\mathbf{w}\in K} \mathbf{w} \cdot \mathbf{L}_t + R(\mathbf{w})$$

Market Making

- Market maker maintains prices $\mathbf{p}_t \in \Pi$ over n contracts
- Contracts are purchased in bundles $\mathbf{r}_t$, and $\mathbf{q}_{t+1} = \mathbf{q}_t + \mathbf{r}_{t+1}$
- Market maker selects prices

$$\mathbf{p}_{t+1} = \arg\max_{\mathbf{p}\in\Pi} \mathbf{p} \cdot \mathbf{q}_t - R(\mathbf{p})$$

Online Linear Opt.

- Learner maintains weights $\mathbf{w}_t \in K$ over n items/experts
- Items have loss vector $\mathbf{l}_t$, cumulatively $\mathbf{L}_{t+1} = \mathbf{L}_t + \mathbf{l}_{t+1}$
- FTRL selects weights

$$\mathbf{w}_{t+1} = \arg\min_{\mathbf{w} \in K} \mathbf{w} \cdot \mathbf{L}_t + R(\mathbf{w})$$

- Learner suffers regret

$$\sum_{t=1}^{T} \mathbf{w}_t \cdot \mathbf{l}_t - \min_{\mathbf{w} \in K} \mathbf{w} \cdot \mathbf{L}_T$$

Market Making

- Market maker maintains prices $\mathbf{p}_t \in \Pi$ over n contracts
- Contracts are purchased in bundles $\mathbf{r}_t$, and $\mathbf{q}_{t+1} = \mathbf{q}_t + \mathbf{r}_{t+1}$
- Market maker selects prices

$$\mathbf{p}_{t+1} = \arg\max_{\mathbf{p} \in \Pi} \mathbf{p} \cdot \mathbf{q}_t - R(\mathbf{p})$$

- MM has worst-case loss

$$\max_{\mathbf{p} \in \Pi} \mathbf{p} \cdot \mathbf{q}_T - \sum_{t=1}^{T} \left(C(\mathbf{q}_t) - C(\mathbf{q}_{t-1}) \right)$$

Online Linear Opt.

- Learner maintains weights $\mathbf{w}_t \in K$ over n items/experts
- Items have loss vector $\mathbf{l}_t$, cumulatively $\mathbf{L}_{t+1} = \mathbf{L}_t + \mathbf{l}_{t+1}$
- FTRL selects weights

$$\mathbf{w}_{t+1} = \arg\min_{\mathbf{w}\in K} \mathbf{w} \cdot \mathbf{L}_t + R(\mathbf{w})$$

- Learner suffers regret

$$\sum_{t=1}^{T} \mathbf{w}_t \cdot \mathbf{l}_t - \min_{\mathbf{w}\in K} \mathbf{w} \cdot \mathbf{L}_T$$

Market Making

- Market maker maintains prices $\mathbf{p}_t \in \Pi$ over n contracts
- Contracts are purchased in bundles $\mathbf{r}_t$, and $\mathbf{q}_{t+1} = \mathbf{q}_t + \mathbf{r}_{t+1}$
- Market maker selects prices

$$\mathbf{p}_{t+1} = \arg\max_{\mathbf{p}\in \Pi} \mathbf{p} \cdot \mathbf{q}_t - R(\mathbf{p})$$

- MM has worst-case loss

$$\max_{\mathbf{p}\in\Pi} \mathbf{p} \cdot \mathbf{q}_T - \sum_{t=1}^{T} \left(C(\mathbf{q}_t) - C(\mathbf{q}_{t-1}) \right)$$

An Example

- Let $K = \Pi = \Delta_n$ and $R(\mathbf{p}) = \sum_i p_i \log p_i$

An Example

- Let $K = \Pi = \Delta_n$ and $R(\mathbf{p}) = \sum_i p_i \log p_i$
- Then there is a closed form solution for the prices/ weights:

$$w_i = \frac{\exp(L_i)}{\Sigma_j \exp(L_j)} \qquad p_i = \frac{\exp(q_i)}{\Sigma_j \exp(q_j)}$$

An Example

- Let $K = \Pi = \Delta_n$ and $R(\mathbf{p}) = \sum_i p_i \log p_i$
- Then there is a closed form solution for the prices/weights:

$$w_i = \frac{\exp(L_i)}{\Sigma_j \exp(L_j)} \qquad p_i = \frac{\exp(q_i)}{\Sigma_j \exp(q_j)}$$

randomized weighted majority / hedge

An Example

- Let $K = \Pi = \Delta_n$ and $R(\mathbf{p}) = \sum_i p_i \log p_i$
- Then there is a closed form solution for the prices/ weights:

$$w_i = \frac{\exp(L_i)}{\Sigma_j \exp(L_j)} \qquad p_i = \frac{\exp(q_i)}{\Sigma_j \exp(q_j)}$$

randomized weighted majority / hedge

logarithmic market scoring rule

More on Choosing *R*

- Interesting market properties can be described in terms of the conjugate…

More on Choosing R

- Interesting market properties can be described in terms of the conjugate…
  - Worst-case market maker loss can be bounded by

$$\sup_{\mathbf{x} \in \mathrm{Hull}(\boldsymbol{\rho})} R(\mathbf{x}) - \inf_{\mathbf{x} \in \mathrm{Hull}(\boldsymbol{\rho})} R(\mathbf{x})$$

More on Choosing R

- Interesting market properties can be described in terms of the conjugate…
  - Worst-case market maker loss can be bounded by

$$\sup_{\mathbf{x} \in \mathrm{Hull}(\boldsymbol{\rho})} R(\mathbf{x}) - \inf_{\mathbf{x} \in \mathrm{Hull}(\boldsymbol{\rho})} R(\mathbf{x})$$

  - Information loss (or the bid-ask spread, or the speed at which prices change) can be bounded too

More on Choosing *R*

- Interesting market properties can be described in terms of the conjugate…
  - Worst-case market maker loss can be bounded by

$$\sup_{\mathbf{x} \in \mathrm{Hull}(\boldsymbol{\rho})} R(\mathbf{x}) - \inf_{\mathbf{x} \in \mathrm{Hull}(\boldsymbol{\rho})} R(\mathbf{x})$$

  - Information loss (or the bid-ask spread, or the speed at which prices change) can be bounded too

Gives us a way to optimize trade-offs in market design!

Example: Permutations

- Suppose our state space is all permutations of *n* items (e.g., candidates in an election, or horses in a race)

Example: Permutations

- Suppose our state space is all permutations of n items (e.g., candidates in an election, or horses in a race)
  - Pair bets: Bets on events of the form "horse i finishes ahead of horse j" for any i, j
  - Subset bets: Bets on events of the form "horse i finishes in position j" for any i, j

Example: Permutations

- Suppose our state space is all permutations of n items (e.g., candidates in an election, or horses in a race)
  - Pair bets: Bets on events of the form "horse i finishes ahead of horse j" for any i, j
  - Subset bets: Bets on events of the form "horse i finishes in position j" for any i, j

- Both known to be #P-hard to price using LMSR [Chen et al., 2008]
- The complex market framework handles both

Example: Permutations

Subset bets ("horse i finishes in position j")

Example: Permutations

Subset bets ("horse *i* finishes in position *j*")

- Hull($\boldsymbol{\rho}$) can be described by a small number of constraints:

$$\sum_j \text{price}(i \text{ in slot } j) = 1 \qquad \sum_i \text{price}(i \text{ in slot } j) = 1$$

Example: Permutations

Subset bets ("horse i finishes in position j")

- Hull($\boldsymbol{\rho}$) can be described by a small number of constraints:

$$\sum_j \text{price}(i \text{ in slot } j) = 1 \qquad \sum_i \text{price}(i \text{ in slot } j) = 1$$

- Easily handled

Example: Permutations

Pair bets ("horse i finishes ahead of horse j")

Example: Permutations

Pair bets ("horse i finishes ahead of horse j")

- Hull($\boldsymbol{\rho}$) is a bit uglier…

Example: Permutations

Pair bets (“horse *i* finishes ahead of horse *j*”)

- Hull(***ρ***) is a bit uglier…
- Solution: Relax the no-arbitrage axiom

Example: Permutations

Pair bets ("horse i finishes ahead of horse j")

- Hull($\boldsymbol{\rho}$) is a bit uglier…
- Solution: Relax the no-arbitrage axiom
  - Allows us to to work with a larger, efficiently specified price space

Example: Permutations

Pair bets ("horse i finishes ahead of horse j")

- Hull($\boldsymbol{\rho}$) is a bit uglier…
- Solution: Relax the no-arbitrage axiom
  - Allows us to to work with a larger, efficiently specified price space
  - But does it increase worst case loss?

Example: Permutations

Pair bets ("horse i finishes ahead of horse j")

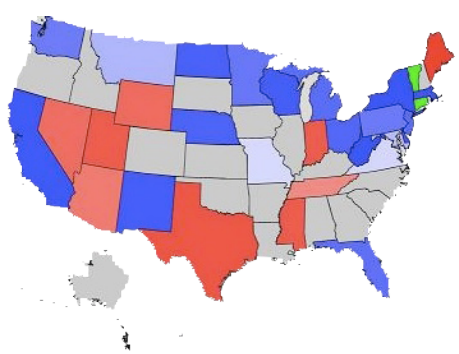
- Hull($\boldsymbol{\rho}$) is a bit uglier…
- Solution: Relax the no-arbitrage axiom
  - Allows us to to work with a larger, efficiently specified price space
  - But does it increase worst case loss? No!

Extensions and Additional Topics

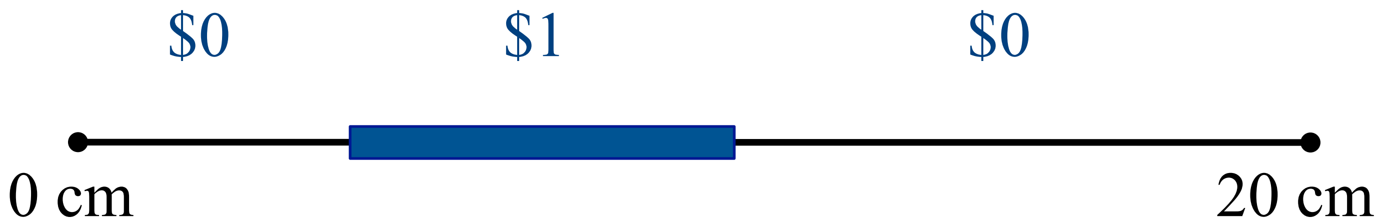
Relaxing No-Arbitrage

- Dudík, Lahaie, and Pennock [2012] pushed on the idea of relaxing no-arbitrage to provide a general constraint generation technique for constructing efficient markets with "approximately consistent" prices

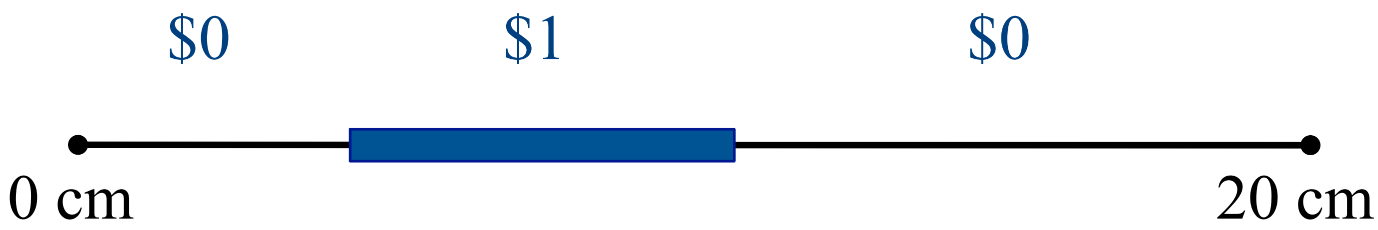
- Used this to implement the WiseQ market which allowed combinatorial bets the 2012 US presidential and senate elections [Dudík et al., 2013]



Continuous Outcome/Contract Spaces

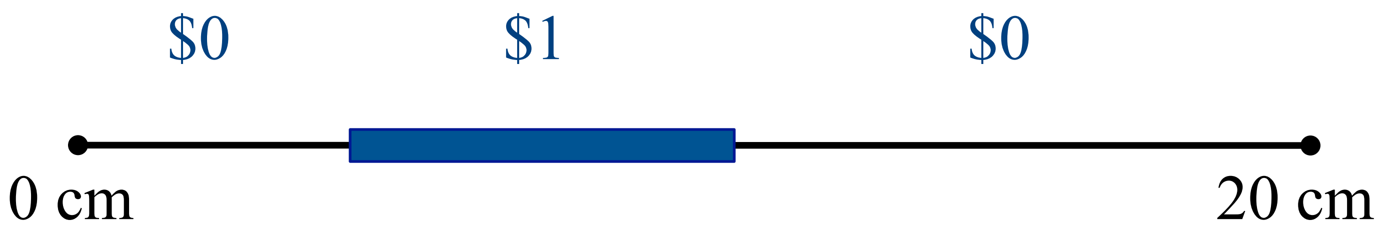


Continuous Outcome/Contract Spaces



- Can discretize the outcome space ex ante, but complexity and worst-case loss grow with the number of outcomes
- Most early attempts to avoid ex ante discretization led to negative results [e.g., Gao and Chen, 2010]
- Chen, Ruberry, and Vaughan [2013] extended the duality framework to markets over continuous outcome spaces, generating markets with bounded worst case loss

Continuous Outcome/Contract Spaces



- Still lots of work to do here quantifying trade-offs between discretization and specially designed markets!!
  - Worst case loss
  - Computational complexity
  - Granularity of predictions

Making a Profit / Adaptive Liquidity

- We have assumed that the market maker is willing to take a potential (bounded) loss in order to obtain information
- The ideas presented here can be modified to yield market makers guaranteed to earn a profit if the volume of trades is sufficiently high and traders disagree [e.g., Othman & Sandholm, 2011; Li & Vaughan, 2013]
- Yields markets with adaptive liquidity
- In the complete market setting, this requires that prices sum to something more than one – adds some ambiguity when backing out probability estimates

Markets & Variational Inference

- The math behind these markets also parallels the math behind variational inference

mean parameter	$\leftrightarrow$	prices
natural parameter	$\leftrightarrow$	quantity vector
sufficient statistics	$\leftrightarrow$	payoff function

- This connection can be used to design new scoring rules [e.g., Lahaie, working paper, 2012]

Price Convergence & Aggregation

- When do security prices converge, and do they reflect the private information or beliefs of the traders?
- Ostrovsky [2012] showed that prices generally converge and incorporate traders' private information if traders are risk neutral and Bayesian with a common prior
- Price convergence also occurs for risk averse traders with heterogeneous beliefs and budgets [Sethi and Vaughan, working paper 2013]
- Lots to do here too! Which model(s) are reasonable? What happens in real markets?

See the tutorial website for additional references:

http://aaaimarketstutorial.pbworks.com